ARI Research Note 97-32

Analysis of the Organization of Lexical Memory

George A. Miller
Princeton University

Research and Advanced Concepts Office
Michael Drillings, Chief

September 1997

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Princeton University

Technical review by

George Lawton

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE 1997, September	2. REPORT TYPE Interim	3. DATES COVERED (from. . . to) March 1990-June 1991
4. TITLE AND SUBTITLE Analysis of the Organization of Lexical Memory		5a. CONTRACT OR GRANT NUMBER MDA903-86-K-0242 5b. PROGRAM ELEMENT NUMBER 0601102A
6. AUTHOR(S) George A. Miller (Princeton University)		5c. PROJECT NUMBER B74F 5d. TASK NUMBER 8522 5e. WORK UNIT NUMBER C26
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Princeton University Princeton, NJ 08544-1010		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-BR 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Research Note 97-32

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

COR: George Lawton

14. ABSTRACT *(Maximum 200 words)*:

The practical outcome of the project, "Analysis of the Organization of Lexical Memory," is an electronic lexical database called WordNet that can be incorporated into computer systems for processing English text. WordNet includes approximately 45,000 lexicalized concepts, providing a coverage equivalent to a handheld dictionary. The database has three components, one each for nouns, verbs, and adjectives. The semantic relations that organize each component are different, but in general a lexicalized concept is represented by a set of synonyms that can be used to express the concept, the familiar semantic relations are represented by labeled pointers between synonyms sets. In order to create the database, programs were written to write and edit lexical files, to convert lexical files into database, to search the database, to strip inflections from search requests, and to display retrieved information for a user. Three user interfaces have been developed for WordNet. (1) The simplest is a commandline version that does not require a windowing system and can run on standard monitors. (2) A browser written for Sun View and for X-11 windows is intended for use with an on-line dictionary; by using WordNet, the dictionary can be searched conceptually as well as alphabetically. (3) A lexical filter written for X-11 windows catches unfamiliar words in a text file and suggests alternative expressions that an author may wish to choose.

15. SUBJECT TERMS

Lexical database   Natural language processing   Lexicography

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	9	

ANALYSIS OF THE ORGANIZATION OF LEXICAL MEMORY

Abstract

The practical outcome of the project, "Analysis of the Organization of Lexical Memory," is an electronic lexical database called WordNet that can be incorporated into computer systems for processing English text. WordNet includes approximately 45,000 lexicalized concepts, providing a coverage equivalent to a handheld dictionary. The database has three components, one each for nouns, verbs, and adjectives. The semantic relations that organize each component are different, but in general a lexicalized concept is represented by a set of synonyms that can be used to express the concept, and familiar semantic relations are represented by labeled pointers between synonym sets. In order to create the database, programs were written to write and edit lexical files, to convert lexical files into a database, to search the database, to strip inflections from search requests, and to display retrieved information for a user.

Three user interfaces have been developed for WordNet. (1) The simplest is a command-line version that does not require a windowing system and can run on standard monitors. (2) A browser written for SunView and for X-11 windows is intended for use with an on-line dictionary; by using WordNet, the dictionary can be searched conceptually as well as alphabetically. (3) A lexical filter written for X-11 windows catches unfamiliar words in a text file and suggests alternative expressions that an author may wish to choose.

Background

The on-line database now known as WordNet began as an experiment designed to test whether certain psycholinguistic claims—namely, that the organization of lexical memory can be represented as a network of labeled nodes (for lexicalized concepts) connected by labeled arcs (for semantic relations between concepts)—could be extended to cover the entire lexical core of English. These claims, which can be referred to generically as the *relational hypothesis*, were stated in the psycholinguistic literature in very general terms, but were usually illustrated with only a handful of carefully chosen lexical items. Moreover, this relational hypothesis contrasted with other psycholinguistic claims, which can be referred to generically as the *componential hypothesis*, to the effect that the organization of lexical memory is best represented by analysis into semantic components, rather than into semantic relations. Fundamental questions about the theory of lexical knowledge—such questions as how much of the descriptive load can be carried by relations and how much by components—were unanswered. In order to pursue such questions, therefore, it was decided to push the relational approach as far as it would go—to apply it literally to the entire substantive lexicon of English—to see where it fails and to discover what kinds of lexical knowledge require more sophisticated analysis.

The experiment can be counted a success, although a relational characterization of lexical memory for all of English could not be implemented as directly as had been anticipated at the beginning; a number of unexpected problems had to be resolved in order to carry it through. An initial decision was made to limit the experiment to semantic relations between open class words; closed class words (prepositions, pronouns, conjunctions, articles, etc.) are better characterized by their syntactic properties and relations, and for practical applications in natural language processing the closed class words should be an integral part of the parsing program. But even for open class words there are differences between parts of speech that a relational representation must respect: for nouns, the relation of class inclusion is most important; for verbs, a complex set of entailment relations is required; and modifiers are best characterized in terms of oppositions. Consequently, discovering what semantic relations to use required three concurrent and related investigations, and resulted in three relatively independent networks: one each for nouns, verbs, and adjectives.

Semantic Relations

What terms should a semantic relation relate? A basic assumption here is that a distinction must be drawn between two common senses of the word "word," between words as concrete

forms (strings of ASCII characters in this instance) and words as abstract concepts that the forms can be used to express. Since computers see character strings where people see concepts, an important goal of this work was to give computers something that could be processed as people process concepts. The initial assumption, therefore, was that semantic relations should be relations between lexicalized concepts.

A wide variety of semantic relations has been described in the technical literature, but few were deemed suitable for this research. The criteria for adoption are simple: (1) Since the basic conception is that of a network, binary (two-term) semantic relations were presupposed. (2) Since broad coverage of the lexicon is a prime consideration, semantic relations with a narrow range of application are neglected (the relation "ancestor of," for example, applies only between kin terms). (3) Since the network is intended for users without special training in linguistics, semantic relations must be intuitively obvious to laypersons. (4) Since workers creating the database are necessarily dependent on standard lexicographic references, semantic relations that are regularly coded in dictionaries and thesauruses are preferred. (5) Since exploration of the network in any direction is desired, only semantic relations that have an obvious reciprocal relation are adopted. A number of semantic relations survived these criteria.

The attempt to limit WordNet to semantic relations between lexicalized concepts failed; in particular, synonymy and antonymy, two basic semantic relations, hold between lexical forms. The other semantic relations, however, are relations between lexicalized concepts.

Synonymy: Two word forms are synonyms if there are linguistic contexts in which one can be substituted for the other without altering the meaning; "snake" and "serpent." (N, V, Adj)

Antonymy: Two word forms are direct antonyms if one is the conventional opposite of the other; "clean" and "dirty." (N, V, Adj)

Hyponymy/Hypernymy: Forms expressing concept A are hyponyms (subordinates, subsets) of forms expressing concept B if A is included in B. If F_A is a hyponym of F_B, then F_B is a hypernym (superordinate, superset) of F_A; "A house is a (kind of) building." (N)

Troponymy: Forms expressing concept A are troponyms of forms expressing concept B if A is a particular manner of doing B; "To march is to walk in a particular manner." The reciprocal relation is also coded in the database, but is called simply "superordinate." (V)

Meronymy/Holonymy: Forms expressing concept A are meronyms of forms expressing concept B if A is a part of B. If F_A is a meronym of F_B, then F_B is a holonym of F_A. Three types of part relations are coded: (1) member ("The navigator is part of the crew"); (2) material ("The paper is part of the page"); (3) component ("The wing is part of the plane"). When the meronym type was uncertain it was coded as a component part. (N)

Entailment: Forms expressing concept A entail forms expressing concept B if the occurrence of B is necessary for the occurrence of A, and F_A and F_B are not related by troponymy; "To fail entails trying." (V)

Cause: A special case of entailment; "To kill is to cause to die." (V)

All of these semantic relations hold between words or concepts in the same syntactic category. Two additional semantic relations—"is an attribute of" and "is a function of"—have not yet been coded. Both require pointers between syntactic categories: between adjectives and nouns in the case of attributes; between verbs and nouns in the case of functions. It is believed that these relations can be added, and that the result will be a better simulation of lexical memory and a more useful database for practical applications.

Although the relations listed above suffice to account for most common word associations, at least one important feature of lexical memory is not captured by a purely relational approach, namely, differences in the familiarity of different words. Although frequency of occurrence is the preferred measure of familiarity, counts broken down by part of speech are not presently available for all of the words included in this database. So an alternative measure was adopted. In general, the more familiar a word is, the more alternative senses it has, so a sense count was made for an on-line dictionary; the results are included in the database for each word by syntactic category.

Finally, since selectional restrictions—the restrictions on noun phrases that can serve as cases (or arguments) of a verb—are so important for syntax, the database includes 33 different sentence frames indicating the admissible syntactic structures for each sense of every verb.

Implementation

In order to realize a computer simulation of this lexical system, it was necessary to have a computer representation for lexicalized concepts as well as lexical forms. The following assumption, therefore, is basic to the implementation: *a lexicalized concept can be represented by a set of word forms that can express that concept when used in appropriate contexts*. For example, the set {case, lawsuit} would represent a different meaning of "case" than would {case, box, carton} or {case, patient}. Such sets of words are called synonym sets or, briefly, *synsets*. Of course, a computer that is given a synset does not "understand" anything, but a human who knows the language will recognize the intended meaning. But the computer should be able to process a synset in a manner analogous to the way people process the corresponding concept.

As work progressed, however, it was discovered that synonyms are not always available to signal conceptual differences between synsets. Therefore, the standard lexicographic method of adding a defining gloss was adopted to clarify the intended distinctions. Since this resort to definitions came relatively late, they are available for only about 30% of the synsets. They are coded parenthetically and can be either displayed or suppressed by the interface.

Given this coding for synonymy, other semantic relations can be coded either by pointers between word forms or by pointers between synsets. For example, the fact that "war" is an antonym of "peace" is coded [war !→ peace], and the fact that tennis is a kind of court game is coded {tennis, lawn_tennis} @→ {court_game}. These semantic relations are entered by lexical coders; the reciprocal relations are then added automatically by a program known as the "grinder," which converts lexical files into a lexical database.

Software developed in order to implement this system is written in C and C++ and includes the following components:

Editor: These programs support the work of entering information into the lexical files. To supplement the editor, there are programs to search and display the contents of on-line dictionaries, to verify the syntax of the lexical files, to recast a noun file in the form of an outline, and to provide an archive to keep track of the files as they are edited and up-dated.

Grinder: This large program turns the lexical files into a database. It first checks for coding errors and requests corrections. Then it inserts all of the reciprocal semantic relations that coders omit, and outputs the result as a coherent database with a unique identifier for every synset. Finally, it constructs an index of the letter strings, listing all of the synsets in which each string appears.

Search routines: A set of routines accepts requests as input and returns information retrieved from the database. A request consists of a letter string and an identifier for the kind of semantic relation that is desired.

Morphology: The WordNet database contains primarily canonical word forms. That is to say, it contains information about the singular "tree" but not about the plural "trees," about present tense "hurl" but not past tense "hurled," etc. For practical applications, therefore, it is necessary to have a morphology program that will transform these inflected forms into the canonical forms contained in the database. This program is fairly conventional. It contains an extensive list of exceptions—words that do not follow the rules of English morphology. If a requested character string is on this list, its canonical form will be used to search the database. If a character string is not on the exception list and is not in the database, the program will attempt to strip inflections from it in order to arrive at a string that can be found in the database. Only if these attempts fail will the program report that the string is not in the database.

Combined with search routines, this morphology program takes inflected inputs and returns canonical outputs, e.g., a request for synonyms of "hurled" will elicit "throw." A more sophisticated morphology program that will return inflected outputs—one that will give "threw" or "thrown" as synonyms of "hurled"—is under development as part of the lexical filter application described below.

Interface: Several interfaces have been created to display information that is retrieved for the user. The simplest is a command-line version that can be used on any monitor. A more elaborate interface, using SunView (a windowing system owned by Sun Microsystems, Inc.) was used for systems development. And an interface using the X-11 window system was developed for general distribution with the database. These interfaces are described in more detail in the section on Applications, below.

Man pages: For Unix systems, a set of man pages is available. A user should look first at wnintro(1), which gives an overview of the man pages in chapter 1 of the manual. They include nverify(1) to describe a program that checks the syntax of lexical files, grind(1) to describe operation of the grinder, wntool(1) for the SunView interface, xwn(1) for the X-11 interface, and wn(1) for the command-line interface. There is also wnintro(5), which introduces wninput(5) for the syntax of the lexical input files and wndb(5) for the syntax of the database itself.

Coverage

The goal for WordNet was to include approximately the same vocabulary that one expects to find in a collegiate dictionary. Because the format is so different from a printed dictionary, however, numerical comparisons cannot be made directly. Three different numbers are needed to characterize the size of WordNet: (1) the number of character strings (ASCII strings); (2) the number of synsets; and (3) the number of unique string-synset combinations. (If the same string occurs in five synsets, it counts as one string but five unique string-synset combinations, i.e., each distinct sense of a string is considered to be a different word.) These numbers, broken down by syntactic category, are given in the following table, where the unique string-synset combinations are referred to simply as "Words."

Category	Strings	Synsets	Words
Nouns	36,114	28,276	48,672
Verbs	9,699	6,087	15,824
Adjectives	12,283	10,620	23,912
Total	58,096	44,983	88,408

Much of the work of creating WordNet, however, consisted of inserting pointers between synsets to represent semantic relations between concepts, and the novelty and utility of the system depends on these relations. The total numbers of pointers for the various semantic relations coded in the database are shown in the following table.

Category	Pointers	Definitions
Nouns	40,087	7,164
Verbs	10,771	2,562
Adjectives	13,854	3,962
Total	64,712	13,688

This table also gives the number of synsets in each syntactic category that have an accompanying parenthetical defining phrase.

Applications

Although initially intended as an experiment, the success of the experiment will be tested by the usefulness of the resulting database. The WordNet database is available for general use in natural language processing and is expected to enrich the content of a variety of practical applications. Three examples were developed under this contract, two of which (a command line interface and a browser) were required in order to develop the database, and one (a lexical filter) is intended to assist writers.

Command line: The simplest interface requires a user to tag the request for information about a word with an indication as to what information is requested. This interface can deal with inflectional morphology. For example, the command line:

```
wn went -synsv
```

returns all synsets for the verb "go." The command with three tags:

```
wn fights -synsn -synsv -synsa
```

will elicit a report for all synsets of "fight" (in this case, as a noun and verb, but not as an adjective). The `wn` command without arguments is a request for help: it produces a list of all the available tags. Definitional glosses will not be shown unless the tag `-d` is inserted immediately following the target word.

Although the command-line interface is simple, some of the commands are relatively complex. For example, the tag `-palln` will not only return the parts that are directly coded as parts of the searchword, but will also list all of the parts that the searchword inherits from its hypernyms.

Browser: The interface used for developing WordNet was called "lexpert" or "browser." Initially, it was a window in the SunView window system; subsequently it was rewritten as an X-11 window. A target word can be typed or dragged to the input slot to start a search. If the word is found in the database, buttons appear indicating that WordNet knows about the word as a noun, or a verb, or an adjective, or some combination. The mouse can then be used to expose a menu that lists all of the kinds of information available about that word. The same searches are available with the browser that are available with the command-line interface, but commands that will not yield information are "greyed out" on the

menu. By selecting from the menu, a user can pursue the particular semantic relation of interest. For nouns, the user may have a choice among synonyms, antonyms, hyponyms, hypernyms, or meronyms, or may ask about the word's familiarity. For verbs, the user may select from synonyms, antonyms, superordinates, troponyms, entailments, cause, familiarity, or sentence frames. For adjectives, the user may select synonyms, antonyms, or familiarity. When this interface is used to write lexical files, it is used in conjunction with on-line dictionaries. Thus it becomes possible to search the dictionary conceptually, not merely alphabetically.

Since inflections are stripped from input requests, the browser can also be used while composing a text file—words in the text can be highlighted with the cursor and dragged to WordNet. The third interface was an attempt to capitalize on this feature.

Filter: The filter program is an attempt to use WordNet as part of a writer's assistant. It is not interactive. It takes a text file as input and goes through it word by word. If a word in the text is not found in WordNet, it is added to a list in a file of "unknown words." Experience with the lexical filter has shown that many of the unknown words are proper nouns, some are typographical mistakes, but some are words that clearly should be added to the WordNet database. If a word in the text is found in WordNet, its familiarity is tested; if it is familiar, the filter does nothing, but if it is unfamiliar, the filter prints out all of the synsets in which the word occurs, accompanying each word with its familiarity value. That is to say, an author is not only told that a word is unfamiliar; an attempt is made to suggest more familiar alternatives.

In its present form, the filter frequently suggests alternatives that are inappropriate. For example, they may be for the wrong part of speech. More often, even when they are in the correct syntactic category, they include other senses of the word. Since the filter responds to unfamiliar words and unfamiliar words are seldom ambiguous, these problems are not severe. But a simple parser (or "parts" program) that could use the context in order to discriminate among nouns, verbs, and adjectives would eliminate syntactic confusions. A more intelligent system would be required to eliminate semantic ambiguity. For example, the text-critiquing program being developed by David Kieras at the University of Michigan is one such intelligent system for assisting writers; Kieras is exploring the use of the semantic information in WordNet to enhance the capabilities of that system. Other opportunities to evaluate WordNet in a testbed provided by a language understanding system are under discussion.

Preliminary results thus confirm the commonsense conclusion that WordNet is best used in conjunction with other components as one part of a more powerful system for natural language processing. The fact that such marriages are possible, however, indicates that WordNet does provide an effective combination of traditional lexicographic information with modern computer technology.

Availability

Copyright to WordNet is held by Princeton University in order to protect the rights of the developers to use their own work and make it available to others, and an application is being filed to protect the term "WordNet." However, an early version has been running on computers at NPRDC, and the database, search code, morphology routines, interface, and man pages (a 7-Mbyte package, WordNet 1.0) are available for public distribution. Inquiries addressed to wordnet@princeton.edu should elicit information about how to obtain these materials via ftp; it is hoped that the Lexical Consortium at New Mexico State University will distribute these materials. If demand justifies it, it can be made available on a cd-rom disk.

Contributors

The following persons, listed in alphabetical order, worked on WordNet prior to July 1991: Amalia Bachman, Richard Beckwith, Marie Bienkowski, Patrick Byrne, Roger Chaffin, George Collier, Michael Colon, Melanie Cook, Christiane Fellbaum, Derek Gross, Brian Gustafson, Philip N. Johnson-Laird, Judith Kegl, Benjamin O. Martin, Elana Messer, George A. Miller, Katherine J. Miller, Antonio Romero, Daniel A. Teibel, Randee Tengi, Anton J. Vishio, Pamela Wakefield.

WordNet Publications

Beckwith, R., Fellbaum, C., Gross, D., and Miller, G. A. (in press). WordNet: A lexical database organized on psycholinguistic principles. In Zernik, U. (ed.), *Using On-line Resources to Build a Lexicon*. Hillsdale, N.J.: Erlbaum.

Beckwith, R., and Miller, G. A. (1990). Implementing a lexical network. *International Journal of Lexicography*, 3, 302-312.

Bienkowski, M. A. (1987). *Tools for Lexicon Construction.* Princeton University Cognitive Science Laboratory, Report No. 10.

Collier, G. H., and Fellbaum, C. (1988). Exploring the verb lexicon with the sensus electronic thesaurus. In *Proceedings of the Fourth Conference of the UW Centre for the New Oxford Dictionary.* Waterloo, Canada: University of Waterloo. Pp. 11-27.

Fellbaum, C. (1990). English verbs as a semantic net. *International Journal of Lexicography*, 3, 278-301.

Fellbaum, C. (in press). Translating with a semantic net: Matching words and concepts. In Lewandowska-Tomascszyk, B. (ed.), *Proceedings of the Lodz Colloquium on Translation and Meaning*. Maastricht, The Netherlands: Euroterm.

Fellbaum, C., and Chaffin, R. (1990). Some principles of the organization of the verb lexicon. 12th Annual Conference of the Cognitive Science Society.

Fellbaum, C., and Kegl, J. (1988). Taxonomic hierarchies in the verb lexicon. Presented at EURALEX Third International Congress, Budapest, Hungary.

Fellbaum, C., and Kegl, J. (1989). Taxonomic structures and cross-category linking in the lexicon. In de Jong, K., and No, Y., (eds.), *Proceedings of the Sixth Eastern States Conference on Linguistics.* Columbus, Ohio: Ohio State University, pp. 93-104.

Fellbaum, C., and Miller, G. A. (1990). Folk psychology or semantic entailment? A reply to Rips and Conrad. *Psychological Review*, 97, 565-570.

Gross, D., Fischer, U., and Miller, G. A. (1989). The organization of adjectival meanings. *Journal of Memory and Language*, 28, 92-106.

Gross, D., and Miller, K. J. (1990). Adjectives in WordNet. *International Journal of Lexicography*, 3, 265-277.

Gustafson, B. (1991). *xwn: An X Windows Interface to the WordNet Lexical Database.* Princeton University Cognitive Science Laboratory, manuscript.

Miller, G. A. (1985). Wordnet: A dictionary browser. *Proceedings of the First Conference of the UW Centre for the New Oxford Dictionary.* Pp. 25-28.

Miller, G. A. (1985). Dictionaries of the mind. *Proceedings, 23rd Annual Meeting of the Association for Computational Linguistics,* University of Chicago. Pp. 305-314. Pp. 277-298.

Miller, G. A. (1986). Dictionaries in the mind. *Language and Cognitive Processes*, 1, 171-185.

Miller, G. A. (ed.) (1990). *Five Papers on WordNet*, special issue of *International Journal of Lexicography*, 3, 235-312.

Miller, G. A. (1990). Nouns in WordNet: A lexical inheritance system. *International Journal of Lexicography*, 3, 245-264.

Miller, G. A. (1991). *The Science of Words.* New York: Scientific American Library.

Miller, G. A. (in press). Lexical echoes of perceptual structure. In *The Perception of Structure*, in honor of W. R. Garner. Washington, D.C.: American Psychological Association.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., and Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography*, 3, 235-244.

Miller, G. A., and Fellbaum, C. (in press). Semantic networks of English. *Cognition.*

Miller, G. A., and Fellbaum, C. (in press). WordNet and the organization of lexical memory. In Swartz, M. (ed.), *The Bridge to International Communication: Intelligent Tutoring Systems for Second Language Learning.* New York: Springer.

Miller, G. A., Fellbaum, C., Kegl, J., and Miller, K. J. (1988). WordNet: An electronic reference system based on theories of lexical memory. *Revue quebecoise de linguistique,* 17, 181-213.

Miller, G. A., Fellbaum, C., Kegl, J., and Miller, K. J. (1988). The Princeton lexicon project: A report on WordNet. In Zigany, J., and Magay, T. (eds.), *Budalex 88: Papers from the Euralex Third International Congress.* Budapest: Akademiai Kiado.

Miller, G. A., and Teibel, D. A. (1991). *A Proposal for Lexical Disambiguation,* 4th DARPA Workshop on Speech and Natural Language, Monterey, California.

Teibel, D. A. (1988). *WordNet User's Guide.* Princeton Univeristy Cognitive Science Laboratory, Report No. 34.

Teibel, D. A. (1988). *A Multilayered Approach to Constructing a Representation of the English Lexicon.* Princeton University Cognitive Science Laboratory, Report No. 35.